View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

ELSEVIER

Chemistry & Biology 8 (2001) 179–187

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Thermodynamic analysis of the interaction between YY1 and the AAV P5 promoter initiator element

Hristo B. Houbaviy [a,1], Stephen K. Burley [a,b,*]

[a] *Laboratory of Molecular Biophysics, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA*
[b] *Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA*

Received 2 August 2000; revisions requested 25 September 2000 revisions received 1 December 2000; accepted 8 December 2000
First published online 3 January 2001

Abstract

Background: We previously determined the co-crystal structure of the zinc finger region of transcription factor YY1 (YY1Δ) bound to the initiator element (Inr) of the adenoassociated virus (AAV) P5 gene promoter [Houbaviy, H.B. et al. (1996) Proc. Natl. Acad. Sci. USA 93, 13577–13582]. Our structure explained both binding specificity and the ability of YY1 to support specific, unidirectional transcription initiation.

Results: To further understand Inr recognition by YY1, we analyzed the YY1Δ–Inr interaction by isothermal titration calorimetry (ITC) and used limited proteolysis, DNase I footprinting and missing nucleoside experiments to show that YY1Δ and full-length YY1 (YY1WT) have indistinguishable DNA binding properties.

Conclusions: YY1 binding occurs at an equilibrium dissociation constant (K_d) of about 1 μM, and exhibits a large negative heat capacity change (ΔC_p). We analyzed the thermodynamic behavior of YY1Δ in terms of buried solvent-accessible surface area resulting from interaction of two rigid bodies, which could not explain our measured value of ΔC_p. We must, therefore, postulate conformational changes in YY1 and/or the Inr or question the validity of current ΔC_p analysis methods for protein–DNA interactions. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Adenoassociated virus; Inr recognition; Thermodynamic analysis; Transcription factor; YY1

1. Introduction

Isothermal titration calorimetry (ITC) has emerged as an essential complement to X-ray crystallographic and solution NMR structural studies of protein–ligand complexes, because it provides a complete description of binding thermodynamics. Any heat capacity change that accompanies binding (ΔC_p) can be measured by ITC. Since ΔC_p is correlated with burial of solvent-accessible surface areas, this technique should, in theory, provide a direct link between thermodynamics and structure.

Physico-chemical studies of the transfer of model compounds between aqueous solution and pure organic phase have been used to derive empirical relationships between ΔC_p and buried solvent-accessible surface area. The following equation, which separates the contributions of buried non-polar (ΔA_{np}) and polar (ΔA_p) surface areas, and the parameterizations of Spolar and co-workers ($\alpha = 0.32 \pm 0.04$; $\beta = -0.14 \pm 0.04$) [2] and Murphy and Freire ($\alpha = 0.45$; $\beta = -0.26$; no error estimates given) [3] were used in this study.

$$\Delta C_p = \alpha \Delta A_{np} + \beta \Delta A_p \quad (1)$$

Eq. 1 is in good agreement with the experimentally determined values of ΔC_p for protein folding, but does not always explain the large negative values typically observed during sequence-specific DNA recognition by proteins [4]. Conformational transitions that accompany DNA binding and result in the burial of additional surface have been invoked to reconcile Eq. 1 with experimental findings (reviewed in [4]). The accuracy of these structural analyses of ΔC_p measurements has been questioned [5,6], and we note that the parameters for Eq. 1 (α, β) obtained by Murphy and Freire are outside of the error bounds given by Spolar and co-workers. Finally, Oas and Toone [7] have argued

[1] Present address: Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

* Correspondence: Stephen K. Burley;
E-mail: burley@rockvax.rockefeller.edu

1074-5521/01/$ – see front matter © 2001 Elsevier Science Ltd. All rights reserved.
PII: S1074-5521(00)00066-1

that the parameterization in Eq. 1 is not valid for processes involving non-protein components, such as highly charged nucleic acids.

Various protein–DNA complexes have been studied by ITC [4,6,8–11]. Here we present the first thermodynamic study of sequence-specific DNA binding by a transcription factor (TF) IIIA-type Cys_2–His_2 zinc finger protein, YY1Δ. Analyses of the behavior of full-length YY1 suggest that our results are relevant to the function of YY1 during transcription initiation from the AAV P5 promoter in vivo.

2. Results and discussion

2.1. *YY1–Inr binding thermodynamics and a structural explanation of ΔC_p*

Representative ITC results for the YY1Δ–Inr complex (Fig. 1A) are provided in Fig. 1B,C and Table 1. At 5°C, binding of YY1Δ to the P5 Inr is endothermic, resulting in positive peaks in the power versus time plot (Fig. 1B). At 25°C, binding is exothermic and the reciprocal behavior is observed (Fig. 1C). The dependence of binding enthalpy (ΔH) on temperature is essentially linear (Fig. 1D), yielding a value of -0.74 (± 0.03) kcal mol^{-1} K^{-1} for ΔC_p. Experimentally determined equilibrium dissociation constants ($K_d = 0.5$–0.8 μM), and derived values of binding free energy (ΔG) do not vary considerably over the temperature range used in this study (Fig. 1D, Table 1). YY1Δ binding is entropically driven at temperatures below 9.7°C (T_h), and enthalpically driven at temperatures above 21.1°C (T_s).

As in most previously published studies, Eq. 1 does not account for the observed ΔC_p of -0.74 kcal mol^{-1} K^{-1} for YY1Δ–Inr complex formation. The theoretical ΔC_p for rigid body association calculated from estimates of ΔA_p and ΔA_{np} derived from our YY1Δ–Inr co-crystal structure (Table 2, Model 1) is -0.16 kcal mol^{-1} K^{-1} if the parameters of Murphy and Freire are used and -0.18 kcal mol^{-1} K^{-1} for the parameterization of Spolar and coworkers (in both cases less than a quarter of the experimentally determined value). To account for the discrepancy between theory and experiment (-0.74 versus

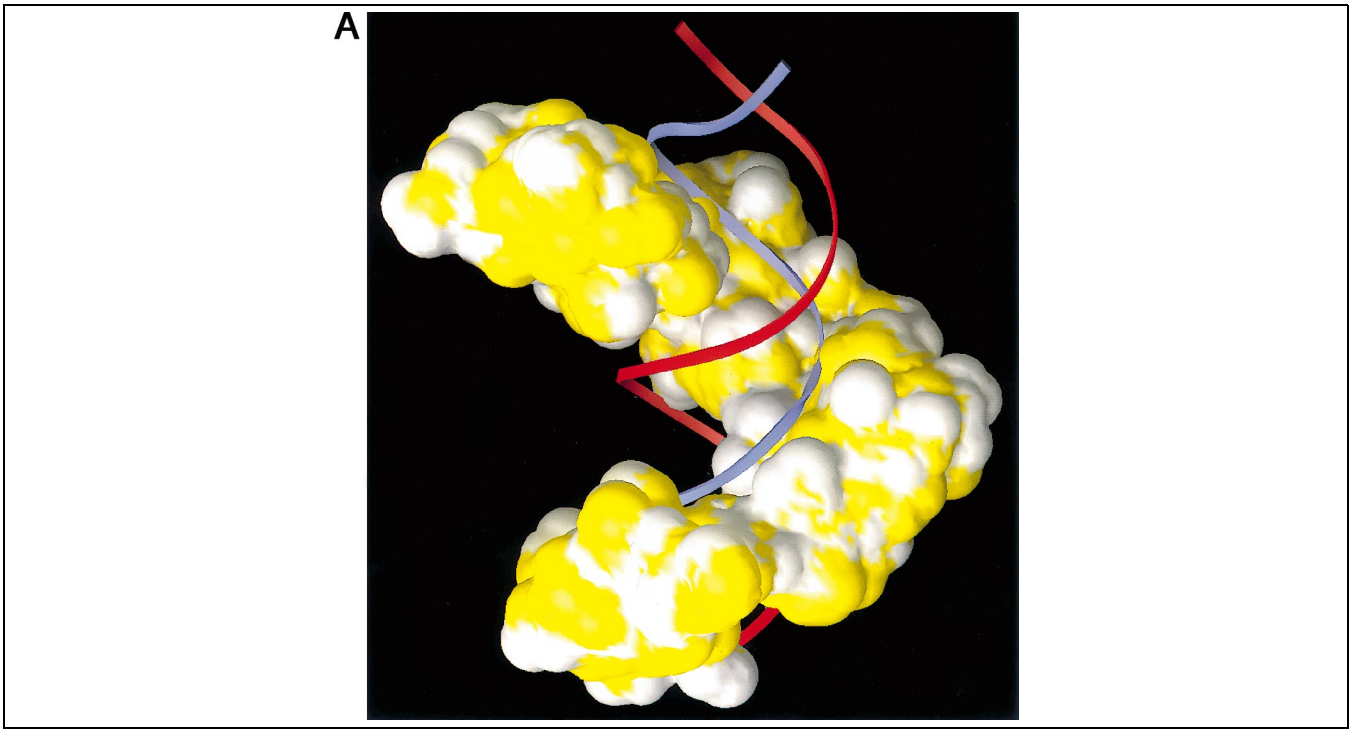

Fig. 1. A: Schematic drawing of the YY1Δ–Inr complex. Protein is depicted as a color coded solvent-accessible surface (polar: yellow; non-polar: white). DNA is shown as a pair of phospho-deoxyribose ribbon backbones, color coded for template (purple) and non-template (red) strands. The DNA-bound form of the protein functions as a contiguous assembly of globular zinc finger domains plus intervening linker regions, which follow the spiral of the widened major groove of the deformed double helix. B,C: Heats absorbed or released during titration of P5 Inr DNA by YY1Δ at 5°C (B) and 25°C (C). The power applied to the microcalorimetric cell (top) and the integrated heats, corrected for dilution effects (bottom) are given as functions of time and the YY1Δ/Inr molar ratio respectively. In B negative peaks are due to dilution effects and constant peak areas at YY1Δ/Inr molar ratios greater than about 1.5 are caused by precipitation of the YY1Δ–Inr complex at 5°C. No precipitation was observed at higher temperatures. D: Diagram representing the thermodynamics of binding of YY1Δ to the Inr as determined by ITC. Values of ΔH (●), $T\Delta S$ (○) and ΔG (□) are plotted as a function of the temperature. The corresponding numerical values appear in Table 1. Dotted lines show the temperatures $T_h = 9.7$°C and $T_s = 21.1$°C at which ΔH and $T\Delta S$ become zero, respectively. For the enthalpy dependence the equation of linear regression fit is $\Delta H = -0.74(\pm 0.03)T + 7.1(\pm 0.7)$; $R^2 = 0.988$; standard deviation of fit = 0.728.

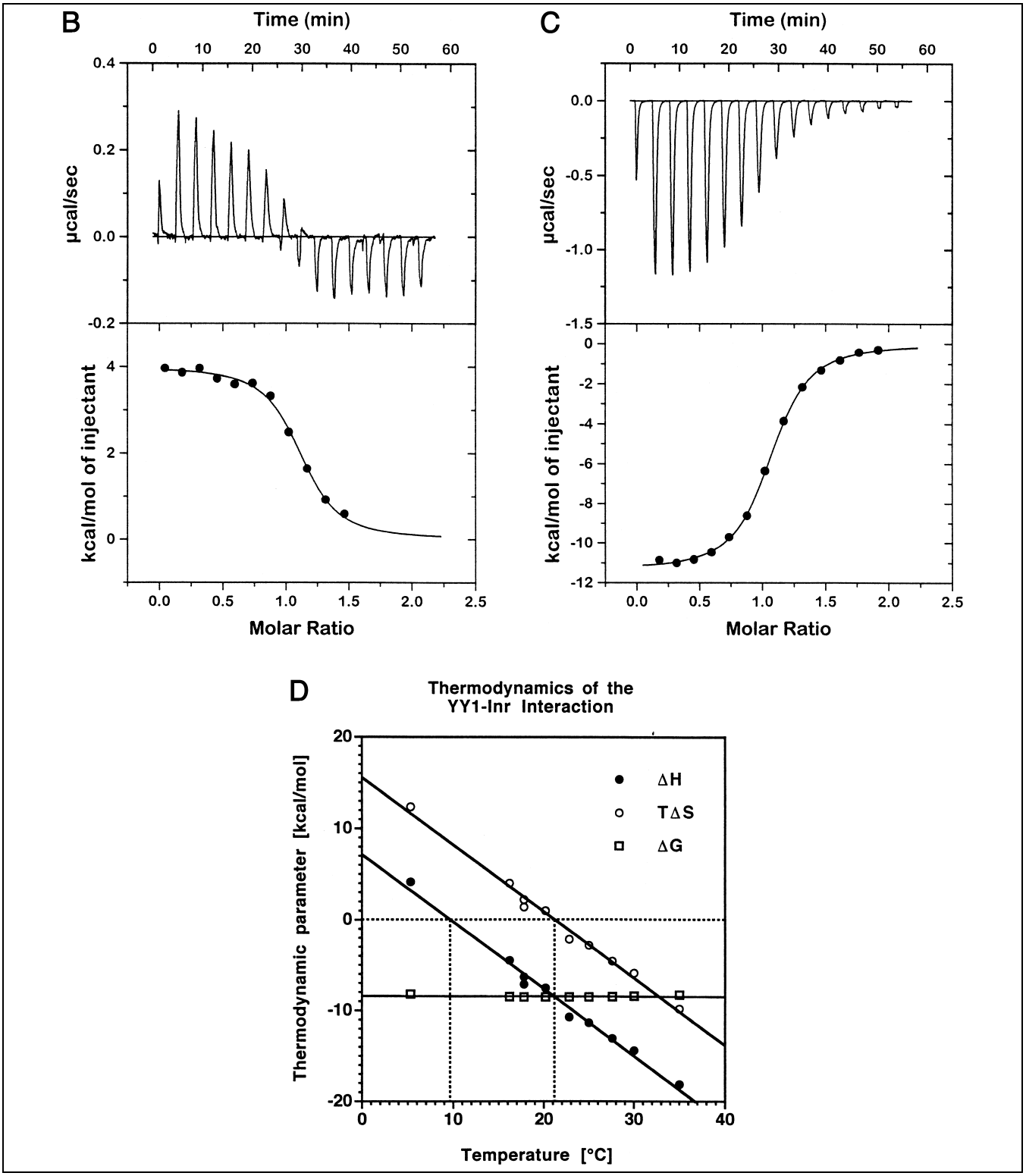

Fig. 1 (*continued*).

-0.18 kcal mol^{-1} K^{-1} at best), we sought conformational transitions within the YY1Δ–Inr complex that result in burial of additional solvent-accessible surface area. Since zinc fingers are well-folded globular domains, conformational changes within the fingers themselves are extremely unlikely. The large negative value observed for ΔC_p could, therefore, result from changes in the conformation of the DNA and/or changes in intramolecular interactions among individual zinc fingers.

First, we consider the possibility that conformational

Table 1
Summary of the YY1–DNA binding thermodynamics

Temperature (°C)	DNA (μmol l^{-1})	YY1 (μmol l^{-1})	Binding stoichiometry	K_d (10^{-6} mol l^{-1})	ΔH (kcal mol^{-1})	ΔG^a (kcal mol^{-1})	$T\Delta S^b$ (kcal mol^{-1})
5.3	24.4	303.1	1.05 ± 0.010	0.51 ± 0.07	4.16 ± 0.06	−8.23	12.39
16.2	24.8	303.1	1.05 ± 0.008	0.47 ± 0.05	−4.51 ± 0.05	−8.52	4.01
17.8	24.8	273.8	1.12 ± 0.006	0.60 ± 0.04	−6.34 ± 0.06	−8.53	2.19
17.8	25.4	273.8	1.20 ± 0.010	0.60 ± 0.07	−7.14 ± 0.08	−8.53	1.39
20.2	24.8	303.1	1.05 ± 0.008	0.49 ± 0.05	−7.57 ± 0.09	−8.55	0.98
22.8	22.9	303.1	1.06 ± 0.007	0.81 ± 0.05	−10.73 ± 0.10	−8.55	−2.19
25.0	24.2	303.1	1.05 ± 0.004	0.56 ± 0.03	−11.38 ± 0.06	−8.53	−2.85
27.6	26.5	303.1	1.00 ± 0.004	0.52 ± 0.03	−13.09 ± 0.08	−8.50	−4.59
30.0	24.1	303.1	1.06 ± 0.007	0.58 ± 0.05	−14.39 ± 0.14	−8.45	−5.94
35.0	23.7	303.1	0.94 ± 0.007	0.80 ± 0.05	−18.13 ± 0.14	−8.31	−9.82

[a]Calculated from the theoretical expression $\Delta G = (T/298)\Delta G_{298} + (1-T/298)\Delta H_{298} - \Delta C_p[298 - T + T\ln(T/298)]$, where T is the temperature in Kelvin, ΔH_{298} is the experimentally determined binding enthalpy at 298 K (25°C) and ΔG_{298} is the free energy of binding at 298 K calculated from the experimentally determined value of K_d. The non-linear regression of data collected at 298 K gave K_d with the lowest relative experimental error (5.4%).
[b]calculated from the fundamental equation $\Delta G = \Delta H - T\Delta S$.

changes occur in the Inr upon binding of YY1Δ. If the Inr assumes a standard B-form structure when not bound to YY1Δ (Table 2, Model 2), ΔA_{np} and ΔA_p for DNA are markedly different from the values calculated for the pure rigid body associations of Model 1. The predicted value of ΔC_p calculated from Eq. 1 is −0.94 to −1.34 kcal mol^{-1} K^{-1} for Model 2, which exceeds the measured value of −0.74 kcal mol^{-1} K^{-1}.

The co-crystal structure of the YY1Δ–Inr complex [1], depicted in Fig. 1A, demonstrates that the four zinc fingers form a contiguous globular protein spiral when bound to the P5 promoter. The three interfaces between pairs of zinc fingers bury portions of each DNA binding domain and their connecting linkers. Analogous intramolecular interfaces were described in the structure of the Zif268–DNA complex [12], and have been extensively analyzed using solution NMR [13,14]. Thus, accurate calculation of ΔA_{np}, ΔA_p, and ultimately ΔC_p for YY1Δ depends on understanding the molecular environments of the three linker regions when the protein is not bound to DNA (Table 2). Making the most generous assumption that the linkers are completely solvent exposed in the absence of a DNA rigid body, the calculated value of ΔC_p would be −0.47 to −0.52 kcal mol^{-1} K^{-1} depending on the parameterization (Table 2, Model 3, Scenario A). Although this result does not account for all of the observed ΔC_p value of −0.74 kcal mol^{-1} K^{-1}, it is remarkable that such a simple model comes so close.

Our thermodynamic data can, therefore, be explained by postulating conformational changes in both YY1 and its target DNA. Rigid body DNA interacting with four rigid body zinc fingers connected by three flexible linkers makes good physico-chemical sense, and accounts for about 70% of the observed ΔC_p value of −0.74 kcal mol^{-1} K^{-1}. The discrepancy of ∼−0.2 kcal mol^{-1} K^{-1} can be made up by postulating DNA conformational changes on protein binding, which is not uncommon for eukaryotic transcription factors (reviewed in [15]). Given the results obtained for Model 2, a surprisingly modest DNA deformation would be required to close the gap between −0.47 to −0.52 and −0.74 kcal mol^{-1} K^{-1}.

At present, it is not possible to determine the relative thermodynamic contributions of conformational changes in the Inr and in YY1Δ during molecular recognition or, indeed, whether such changes occur at all. Shi and Berg [16] have analyzed changes in topoisomer distributions of supercoiled circular DNA upon specific binding of Zif268 and another TFIIIA-like transcription factor, Sp1. Their data provide direct experimental support for a model in which DNA conformational changes are induced by Cys_2–His_2 zinc finger protein binding. Our thermodynamic results suggest that the AAV P5 Inr may indeed differ from standard B-form DNA prior to binding of YY1, which triggers further conformational changes. Mutations in the linker regions of zinc finger proteins are known to have considerable effects on the DNA binding affinity [17,18]. Indeed, alteration of linker residues that do not make contacts with the DNA in the TFIIIA–DNA complex [13,14,19] can abolish or significantly impair DNA binding [20–22]. The above observations taken together with our own ITC data make a dual (both DNA and protein) conformational change model plausible.

We must also consider the possibility that the discrepancy between observed and calculated ΔC_p values is due to flaws in the theoretical analysis of our experimental results. ΔC_p is a parameter intimately linked to the reorganization of hydration shells upon association. Thus, functional group contributions to ΔC_p may not be additive and would depend strongly on the local chemical environment (reviewed in [7]). This argument is particularly apropos for YY1 given the large number of well-ordered water molecules revealed by the YY1Δ–Inr co-crystal structure. Many of these hydration waters are bound to the narrowed minor groove of the Inr DNA, a phenomenon that has no counterpart in protein physical chemistry. Simply including well-ordered solvent into the surface area calculations

as if it were an intrinsic part of the complex, would be as much of an oversimplification of the physical chemistry of hydration as not considering hydration water at all. Thus, the parameterizations of Spolar and co-workers and Freire and Murphy, which work relatively well in predicting protein folding parameters may not be applicable to protein–DNA recognition. Indeed, these parameters appear to be quite unsatisfactory for interactions involving non-protein components [7].

Resolving the discrepancy between the measured and calculated ΔC_p of YY1Δ–Inr complex formation would undoubtedly benefit from knowledge of the structures of the free Inr and YY1Δ. We must, however, point out that even where high-resolution structures of both the DNA–protein complex and the unbound components are available, such as the recognition of the tryptophan operator by the tryptophan repressor, Eq. 1 did not accurately predict the measured value of ΔC_p [6]. Ladbury and co-workers have argued that this discrepancy is due to the tightening of vibrational modes at the interface between protein and DNA in the complex, an elaboration which is not embodied by Eq. 1 [6]. Inaccurate estimates of the coefficients (α, β) in Eq. 1, are an equally likely explanation. Clearly, a better understanding of the thermodynamics of protein–DNA association and a better parameterization of ΔC_p will only emerge as more experimental data on protein–DNA complexes with known structures are accumulated.

2.2. *The P5 Inr binding specificities of YY1Δ and YY1WT are identical*

ITC demonstrates a micromolar equilibrium dissociation constant (K_d) for the YY1Δ–Inr complex (Table 1), which is significantly weaker than the nanomolar K_d values observed for many protein–DNA complexes. Unlike electrophoretic gel mobility shift assays, quantitative footprinting, and filter binding studies (all three of which involve certain implicit assumptions about the stability of the complex under investigation) ITC provides for true equilibrium measurements of K_d. The P5 Inr oligonucleotide used in this study is identical to the DNA fragment bound to YY1Δ in our YY1Δ–Inr co-crystal structure [1]. This sequence represents the natural P5+1 YY1 binding site, which has been extensively characterized both in vivo and in vitro [23–26]. We are confident, therefore, that the relatively weak K_d obtained with ITC is not due to the use of a non-physiological binding site, nor is it caused by experimental limitations of a non-equilibrium measurement technique. Are the ITC results presented above relevant to the behavior of full-length YY1 (YY1WT) in vivo?

Because the full-length protein is aggregation prone at higher concentrations, ITC studies of YY1WT were not feasible. Apparent K_ds for the YY1WT–Inr and the YY1Δ–Inr complexes are equivalent in gel mobility shift assays [1], suggesting that our thermodynamic analysis of the YY1Δ–Inr complex provides useful data pertaining to P5 Inr recognition by YY1WT. To further substantiate this conclusion, we compared the YY1Δ–Inr and the YY1WT–Inr complexes using limited proteolysis, DNase I footprinting, and missing nucleoside experiments.

Subtilisin quantitatively cleaves YY1WT within 20 min, yielding a limit digest of about 14.7 kDa. The reaction time course is not affected by the presence of the P5 Inr (Fig. 2A,B). Analyses of the reaction products by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) revealed two transient species predominant at $t = 1$ min, with molecular masses of $17\,600 \pm 50$ and $17\,900 \pm 50$, and two relatively stable polypeptides present

Table 2
Calculation of ΔC_p due to solvent-accessible surface area burial

	Solvent-accessible surface area contributions			ΔC_p (kcal mol^{-1} K^{-1})[d]
	Total (Å^2)	Non-polar (Å^2)	Polar (Å^2)	
Model 1: Rigid body YY1–DNA association				
YY1–DNA complex	13 184	5 939	7 244	
DNA from complex	7 376	2 046	5 329	
YY1 from complex	8 550	5 122	3 428	
Net change	−2 742	−1 229	−1 513	−0.18/−0.16
Model 2: YY1–DNA association assuming conformational changes in DNA and rigid YY1				
YY1–DNA complex	13 184	5 939	7 244	
free DNA B-form	7 089	3 607	3 482	
YY1 from complex	8 550	5 122	3 428	
net change	−2 455	−2 789	334	−0.94/−1.34
Model 3: Net changes due to zinc finger packing				
Scenario A[a]	−1 980	−1 229	−751	−0.29/−0.36
Scenario B[b]	−1 347	−908	−438	−0.23/−0.29
Scenario C[c]	−1 083	−677	−406	−0.16/−0.20

[a] Linkers not stably associated with any of the zinc fingers in free YY1.
[b] Linkers packed against their N-terminal zinc finger in free YY1.
[c] Linkers packed against their C-terminal zinc finger in free YY1.
[d] According to Spolar and co-workers [2]/Murphy and Freire [3].

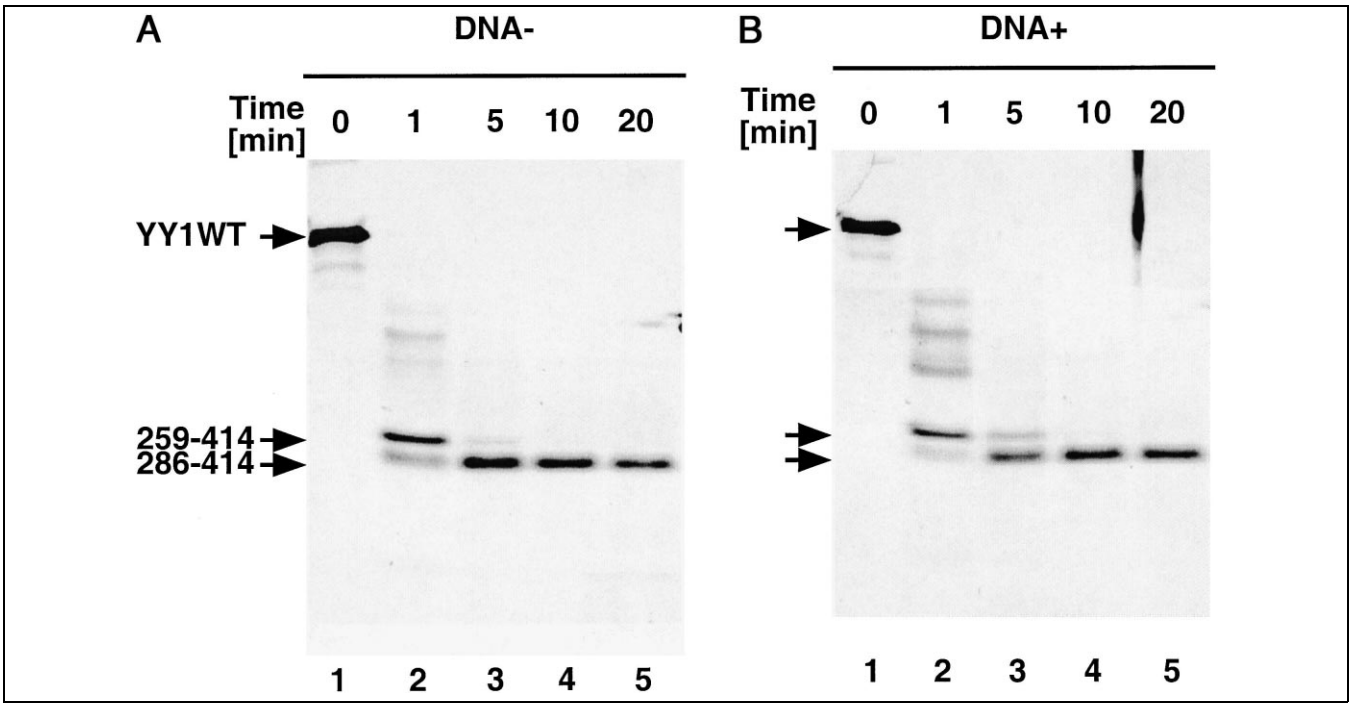

Fig. 2. A,B: Proteolysis of YY1WT by subtilisin in the absence (A) or presence (B) of P5 Inr DNA, monitored by Tris–tricine SDS–gel electrophoresis. The times at which aliquots were removed for analysis are shown above each lane. The initial compositions of the reaction mixtures (0 min) are given in lane 1. The positions of YY1WT and the two stable proteolytic fragments (residues 259–414 and residues 286–414) are indicated by arrows.

at all reaction times, with molecular masses of 15 300 ± 50 and 14 700 ± 50 (data not shown). Electrospray ionization mass spectrometry (ESI-MS) provided higher accuracy masses of 14 697 ± 2, 15 368 ± 3, and 17 613 ± 4 for three of the four fragments detected by MALDI-MS (data not shown). With the ESI-MS data and the results of amino-terminal protein sequencing, we were able to conclusively identify the 17.6 kDa (residues 259–414) and 14.7 kDa (residues 268–414) fragments (Fig. 2A,B). Thus, our proteolysis data suggest that there is little or no globular structure in YY1WT outside the confines of the four zinc fingers. The proteolytically stable 14.7 kDa fragment (residues 286–414) includes only the four zinc fingers. This limit digest pattern is not affected by the presence of P5 Inr, demonstrating that the zinc finger region (YY1Δ) suffices for binding of YY1 to the P5 initiator element. Similar proteolysis studies with YY1Δ revealed no significant cleavage products (data not shown).

DNase I cleavage of the P5 Inr reveals identical footprints for YY1WT and YY1Δ between positions −15 and +15 base pairs flanking the major transcription start site (Fig. 3A). To compare the protein–DNA interfaces formed by YY1Δ and YY1WT at higher resolution, missing nucleoside experiments (Section 4) were performed. The results shown in Fig. 3B,C,D,E demonstrate that YY1WT and YY1Δ exhibit identical binding to the P5 initiator element. On the non-template strand, both YY1WT and YY1Δ require bases in the region extending from −10 to +1, whereas on the template strand the region between +7 and −5 is important for productive interactions. These findings are entirely consistent with our co-crystal structure of the YY1Δ Inr complex (see Fig. 2b in [1]).

Taken together, these comparative results suggest that our thermodynamic analysis of YY1Δ is relevant to the behavior of full-length YY1. The relatively weak K_d for specific YY1Δ–DNA complex formation measured using ITC is compatible with the observed behavior of YY1WT in vitro. During transcription initiation from the AAV P5 promoter by RNA polymerase II (pol II), YY1 forms a multiprotein DNA complex consisting of the P5 Inr, YY1, TFIIB and pol II [27]. TFIIB, and possibly pol II, could further stabilize the YY1–DNA complex, as is the case when TFIIB recognizes the preformed complex of TATA box binding protein and the TATA element of the adenovirus major late promoter (reviewed in [28]).

3. Significance

In eukaryotic class II promoters the TATA box and the pyrimidine-rich Inr elements act independently or synergistically to direct the assembly of the RNA polymerase II pre-initiation complex. Our previous determination of the co-crystal structure of the YY1Δ–Inr complex provided the first atomic resolution glimpse of a eukaryotic transcription factor recognizing an Inr element, yielded mechanistic insight into Inr function, and explained the specificity of the YY1–Inr interaction.

Here we supplement our structural studies with isother-

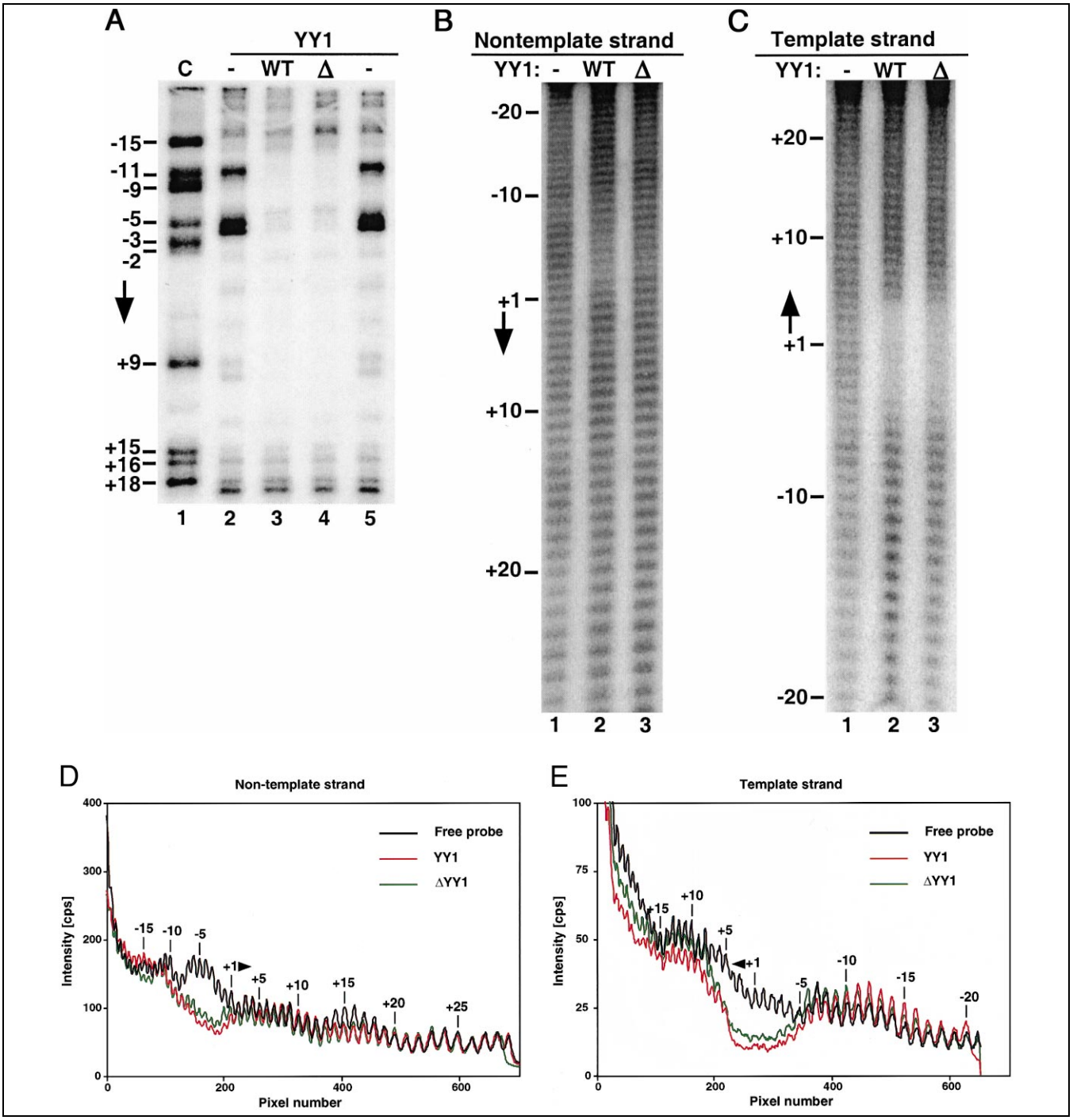

Fig. 3. A: Comparison of the DNase I footprints of YY1WT (lane 3) and YY1Δ (lane 4) on the P5 Inr non-template strand. Digests of the free probe are shown in lanes 2 and 5. The identities of the DNase I cleavage products were determined from the cytosine-specific Maxam–Gilbert sequencing reaction shown in lane 1. Positions relative to the major transcription start site and the direction of transcription (arrow) are given on the left. B,C: Missing nucleoside analysis of the binding of YY1WT (lane 2) and YY1Δ (lane 3) to the P5 Inr non-template (B) and template (C) strands. Nucleotide positions relative to the major transcription start site are indicated on the left and were deduced by comparison with Maxam–Gilbert sequencing reactions (not shown). Arrows indicate the direction of transcription. D,E: Quantitations of the autoradiograms shown in B and C respectively (free probe: black, YY1WT: red, YY1Δ: green). Peaks are labeled according to their relative positions with respect to the major start site and the direction of transcription is indicated by arrows.

mal titration calorimetry measurements, which provide a complete description of the binding thermodynamics. Our physico-chemical analysis improves understanding of the function of YY1 at the AAV P5 Inr, adds to a growing body of ITC measurements of DNA–protein complex formation, and provides valuable general insights into DNA recognition by TFIIIA-type zinc finger proteins.

Binding of YY1Δ to the Inr results in a large negative

ΔC_p, which cannot be explained by a simple rigid body protein–DNA interaction. Burial of linker amino acid residues between adjacent YY1 zinc fingers, however, accounts for much of the discrepancy between observed and predicted ΔC_p and further modest conformational changes within the Inr would suffice to explain the observed ΔC_p. While postulating such conformational changes is tempting, the discrepancy between observed and predicted ΔC_p also strengthens the argument for re-evaluation of the relationship between ΔC_p of complex formation and the three-dimensional structures of protein–DNA complexes.

Unlike other methods ITC gives a reliable measurement of the binding affinity. Thus, the surprisingly weak dissociation constant of the YY1Δ–Inr complex in vitro and our comparative analysis of Inr recognition by YY1WT suggest that additional factors, such as TFIIB and RNA polymerase II, stabilize the complex in vivo.

4. Materials and methods

4.1. Preparation of proteins and nucleic acids

Synthesis and purification of the Inr oligonucleotide have been described previously [1]. YY1WT and YY1Δ were expressed as insoluble proteins in *Escherichia coli* using published methods [1]. Protein and nucleic acid concentrations were determined by quantitative amino acid analyses and UV absorbance spectroscopy, respectively.

4.2. Isothermal titration calorimetry

Binding enthalpies were measured with a Microcal isothermal titration calorimeter. YY1Δ and Inr DNA were dialyzed against a common reservoir containing 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 100 μM zinc acetate, and 1 mM Tris–(2-carboxyethyl) phosphine–HCl (Sigma Chemical Co.). Non-linear regression of the experimental data was performed with Origin (Microcal Software Inc.). Solvent-accessible surface areas were calculated with the method of Lee and Richards [29] using a water probe radius of 1.4 Å.

4.3. Limited proteolysis and mass spectrometry

Digestion with subtilisin was performed in 20 mM HEPES pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 50 μM zinc acetate, and 1 mM Tris–(2-carboxyethyl) phosphine–HCl (Sigma Chemical Co.). The concentration of YY1WT was between 0.5 and 1.0 mg/ml and the protein to protease ratio was 100:1 (w/w). For MALDI-MS analyses, reaction aliquots were mixed in a ratio of 1:5 (v/v) with a saturated α-cyano-4-hydroxy cinnamic acid matrix solution (Sigma Chemical Co.), and prepared in a 1:3:2 (v/v/v) part mixture of formic acid, water and 2-propanol. Tris–tricine SDS–gel electrophoresis was performed according to [30].

4.4. DNase I footprinting and missing nucleoside experiments

DNA probes were prepared by digesting plasmid pP5+1 [24] with *Eco*RV and *Hin*dIII (non-template strand probe) or *Bgl*II and *Pst*I (template strand probe). The desired DNA fragments were isolated by native polyacrylamide gel electrophoresis and singly end labeled with modified phage T7 DNA polymerase (Sequenase version 2.0; Amersham). Missing nucleoside experiments were performed as described in [31]. For DNase I footprinting, YY1WT or YY1Δ and 300 fmol probe were incubated in 8% glycerol, 20 mM HEPES pH 7.5, 60 mM NaCl, 5 mM dithiothreitol, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 100 μg/ml bovine serum albumin, and 50 μg/ml poly[dI:dC] for 30 min. DNA was digested at room temperature for 1 min by adding 0.004 U DNase I (Worthington) to each binding reaction.

Acknowledgements

We thank Dr. Natalia Rodionova for help with the ITC experiments, and Drs. K.S. Gajiwala, J. Kuriyan, G.A. Petsko, R.G. Roeder, T. Shenk, and A. Usheva for many useful discussions. S.K.B. is an investigator in the Howard Hughes Medical Institute, and H.B.H. was supported as a Rockefeller University Graduate Fellow.

References

[1] H.B. Houbaviy, A.A. Usheva, T. Shenk, S.K. Burley, Co-crystal structure of YY1 bound to the Adeno-associated virus P5 Initiator, Proc. Natl. Acad. Sci. USA 24 (1996) 13577–13582.
[2] R.S. Spolar, J.R. Livingstone, M.T. Record Jr., Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water, Biochemistry 31 (1992) 3947–3955.
[3] K.P. Murphy, E. Freire, Thermodynamics of structural stability and cooperative folding behaviour in proteins, Adv. Protein Chem. 43 (1992) 313–361.
[4] R.S. Spolar, T.M. Record, Coupling of local folding to site-specific binding of proteins to DNA, Science 263 (1994) 777–784.
[5] J. Janin, Angstroms and calories, Structure 5 (1997) 473–479.
[6] J.E. Ladbury, J.G. Wright, J.M. Sturtevant, P.B. Sigler, A thermodynamic study of the trp repressor-operator interaction, J. Mol. Biol. 238 (1994) 669–681.
[7] T.G. Oas, E.J. Toone, Thermodynamic solvent isotope effects and molecular hydrophobicity, Adv. Biophys. Chem. 6 (1997) 1–52.
[8] M. Oda, K. Furukawa, K. Ogata, A. Sarai, H. Nakamura, Thermodynamics of specific and non-specific DNA binding by the c-Myb DNA-binding domain, J. Mol. Biol. 276 (1998) 571–590.
[9] C. Berger, I. Jelesarov, H.R. Bosshard, Coupled folding and site-specific binding of the GCN4-bZIP transcription factor to the AP-1 and ATF/CREB DNA sites studied by microcalorimetry, Biochemistry 35 (1996) 14984–14991.
[10] T. Lundback, T. Hard, Sequence-specific DNA-binding dominated by dehydration, Proc. Natl. Acad. Sci. USA 93 (1996) 4754–4759.
[11] A.G.E. Kunne, M. Sieber, D. Meierhans, R.K. Allemann, Thermodynamics of the DNA binding reaction of transcription factor MASH-1, Biochemistry 37 (1998) 4217–4223.
[12] N.P. Pavletich, C.O. Pabo, Zinc finger-DNA recognition: crystal

structure of a Zif268-DNA complex at 2.1 Å, Science 252 (1991) 809–817.

[13] M.P. Foster et al., Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA, Nat. Struct. Biol. 8 (1997) 605–608.

[14] D.S. Wuttke, M.P. Foster, D.A. Case, J.M. Gottesfeld, P.E. Wright, Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity, J. Mol. Biol. 273 (1997) 183–206.

[15] G. Patikoglou, S.K. Burley, Eukaryotic transcription factor-DNA complexes, Annu. Rev. Biophys. Biomol. Struct. 26 (1997) 289–325.

[16] Y. Shi, J.M. Berg, DNA unwinding induced by zinc finger protein binding, Biochemistry 35 (1996) 3845–3848.

[17] S.K. Thukral, M.L. Morrison, E.T. Young, Alanine scanning site-directed mutagenesis of the zinc fingers of transcription factor ADR1: residues that contact DNA and that transactivate, Proc. Natl. Acad. Sci. USA 88 (1991) 9188–9192.

[18] T.E. Wilson et al., In vivo mutational analysis of the NGFI-A zinc fingers, J. Biol. Chem. 267 (1992) 3718–3724.

[19] R.T. Nolte, R.M. Conlin, S.C. Harrison, R.S. Brown, Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex, Proc. Natl. Acad. Sci. USA 95 (1998) 2928–2943.

[20] Y. Choo, A. Klug, A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA, Nucleic Acids Res. 21 (1993) 3341–3346.

[21] K.R. Clemens et al., Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein, Science 260 (1993) 530–533.

[22] K.R. Clemens et al., Relative contributions of the zinc fingers of transcription factor IIIA to the energetics of DNA binding, J. Mol. Biol. 244 (1994) 22–35.

[23] R. Hyde-DeRuyscher, E. Jennings, T. Shenk, DNA binding sites for the transcriptional activator/repressor YY1, Nucleic Acids Res. 23 (1995) 4457–4465.

[24] E. Seto, Y. Shi, T. Shenk, YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro, Nature 354 (1991) 241–245.

[25] Y. Shi, E. Seto, L.-S. Chang, T. Shenk, Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A, Cell 67 (1991) 377–388.

[26] S.R. Yant et al., High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human beta globin cluster, Nucleic Acids Res. 23 (1995) 4353–4362.

[27] A. Usheva, T. Shenk, TFIIB and the large subunit of RNA polymerase II bind directly to YY1 to mediate transcription initiation, Proc. Natl. Acad. Sci. USA 93 (1996) 13571–13576.

[28] D.B. Nikolov et al., Crystal structure of a TFIIB-TBP-TATA element ternary complex, Nature 377 (1995) 119–128.

[29] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, J. Mol. Biol. 55 (1971) 379–400.

[30] H. Schagger, G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, Anal. Biochem. 166 (1987) 368–379.

[31] W.J. Dixon et al., Hydroxyl radical footprinting, Methods Enzymol. 208 (1991) 380–413.